

SEQUENCE LISTING

<110> XU, Ming-Qun
 EVANS, Thomas C.

<120> Intein Mediated Peptide Ligation

<130> NEB-150PUS

<140> 09/786,009

<141> 1999-09-30

<150> 00/102,413

<151> 1999-09-30

<150> ECT/US99/22776

<151> 1999-09-30

<160> 9

<170> PatentIn Ver. 2.0

<210> 1

<211> 43

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: the modified
 C-terminal splice junction of the intein from the
 gyrA gene of Mycobacterium xenopi

<400> 1

gggttcgtcag ccacgtact ggctcaccg gttgatagc gca

43

<210> 2

<211> 19

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: the
 complementary strand of the C-terminal splice
 junction of the modified intein from the gyrA
 gene of Mycobacterium xenopi

<400> 2

ggtatcattc ggtaggaca ctatcttgcg tgcctga

<210> 3

<211> 69

<212> DNA

<213> Artificial Sequence

4200 -

4203 - Description of Artificial Sequence: the polylinker
sequence inserted upstream of the modified intein
from the gyrA gene of Mycobacterium xenopi

4400 - 3

tcgaattctag acatatggcc atgggtggcg gcgcgcctcga ggcctcttcc tgcattcacgg 60
cagatgca 68

4210 - 4

4211 - 69

4212 - DNA

4213 - Artificial Sequence

4214 -

4215 - At position 41, "H" = A or C or T.

4220 -

4223 - Description of Artificial Sequence: the
complementary strand of the polylinker inserted
upstream of the modified intein from the gyrA
gene of Mycobacterium xenopi

4400 - 4

cctggcgcctcga ggcctcttcc tgcattcacgg 60
cagatgca 68

4210 - 5

4211 - 6109

4212 - DNA

4213 - Artificial Sequence

4214 -

4223 - Description of Artificial Sequence: pTXB1 plasmid
sequence containing the modified intein from the
gyrA gene of Mycobacterium xenopi

4400 - 5

aaattatgta ggtggcaact ttgggggaaa tgtggcgogga aacctatttt gtattatttt 60
aaattatgat tcaaatatgt atcggctcat gagacaataa ccccgataaa tgcattcaata 120
aaattatgta tcaaatatgt atcggctcat gagacaataa ccccgataaa tgcattcaata 180
tggcgacatt tgccttcctg tttttgctca cccagaaaag atgggtgaaag taaaagatg 240
tggcgacatt tgccttcctg tttttgctca cccagaaaag atgggtgaaag taaaagatg 300
aaattatgta tcaaatatgt atcggctcat gagacaataa ccccgataaa tgcattcaata 360
aaattatgta tcaaatatgt atcggctcat gagacaataa ccccgataaa tgcattcaata 420
aaattatgta tcaaatatgt atcggctcat gagacaataa ccccgataaa tgcattcaata 480
aaattatgta tcaaatatgt atcggctcat gagacaataa ccccgataaa tgcattcaata 540
aaattatgta tcaaatatgt atcggctcat gagacaataa ccccgataaa tgcattcaata 600
aaattatgta tcaaatatgt atcggctcat gagacaataa ccccgataaa tgcattcaata 660
aaattatgta tcaaatatgt atcggctcat gagacaataa ccccgataaa tgcattcaata 720
aaattatgta tcaaatatgt atcggctcat gagacaataa ccccgataaa tgcattcaata 780
aaattatgta tcaaatatgt atcggctcat gagacaataa ccccgataaa tgcattcaata 840
aaattatgta tcaaatatgt atcggctcat gagacaataa ccccgataaa tgcattcaata 900
aaattatgta tcaaatatgt atcggctcat gagacaataa ccccgataaa tgcattcaata 960
aaattatgta tcaaatatgt atcggctcat gagacaataa ccccgataaa tgcattcaata 1020
aaattatgta tcaaatatgt atcggctcat gagacaataa ccccgataaa tgcattcaata 1080

ttttgtataag	caaatatttta	aattgtaaac	gttaatatatt	tgttaaaaatt	cgogttaaat	1140
ttttgttaaaa	tcagctcatt	ttttaaccaa	taggcgcgaaa	toggcacaat	ccrtttataaa	1200
tcaaaaagaat	agcccgagat	agggltgagt	gttgttccag	tcgggacaaa	gagtccatta	1260
ttaaagaaag	tggactccaa	cgtccaaaggy	cgaaaaacgg	tcctatccagg	cgatggccca	1320
ctacgtgaa	cattcaccbaa	atcaagtttt	ctgggggtoga	ggtgctgttaa	agcactaaat	1380
cgaaacccca	aaggggagccc	ccgattttaga	gottgacggg	gaaagccggc	gaaogtggcg	1440
agaaaggaa	ggaagaaaag	gaaaggagcg	ggcgccagg	cgctggcaag	tgtagcggtc	1500
acgtcgcgcg	taaccaccca	acccgcccgg	cttaattggc	cgctacagg	cgcgtaaaag	1560
gacttaggtg	aagatccctt	ttgataatct	cattgacaaa	atcccttaac	gtgagttttc	1620
gtccactga	ggtccagacc	ccgtagaaaa	gatcaaggga	tctctcttgag	atcccttttt	1680
ctcgcgcgta	atctgctgct	tgcatacaaa	aaaaccacgg	ctaccagcgg	tggttctgtt	1740
gcgggatcaa	gagctaccba	ctctctctcc	gaaggtaact	ggcttcagca	gagcgcgat	1800
accaaataat	gtctctttag	tgtagccgta	gttaggcba	cacttcaaga	actctgtago	1860
acggccataa	taactctgctc	tgttaatctc	gtatccagtg	gctgctgca	ctggcgataa	1920
gcgtgtctt	acggggttgg	actcaagag	atagttacc	gataaggcgc	agcggtcggg	1980
ctgaacgggg	ggtctgtgca	cacagcccag	cttggaagca	acgaactaca	ccgaactgag	2040
atacctacag	cgtgagctat	gagaaaagcg	caagcttccc	gaaggggagaa	aggcggaacg	2100
gtatccggta	agcggcaggg	tgggaacagg	agagggcag	agggaagctc	cagggggaaa	2160
cgctcggtat	cttctatagt	ctgctcgggt	tggccacttc	tgaacttgag	gtgacttttc	2220
gtgatgctcg	tcaggggggg	ggagccctat	gaaaaacgcc	agcaacggcg	ctcttttaag	2280
gttctctggc	ttttgctggc	ctctctgctc	cattgtctct	cttgcgcttat	ccctcgattc	2340
tgtggataac	cgtattacc	ctcttgagtg	agctgatacc	gctccggcca	gcggaaagac	2400
cgagcccaat	gagtccgtga	ggagggagc	tatgtggaat	tctccattaca	atctgctctg	2460
atgcacaaa	gcttaagccag	taaacactcc	gctatcgcta	cgtgactggg	tcattggtcg	2520
gcacgggaa	caactaacba	ccctcgaggg	gctcgagaa	gctggcgctg	tcacgggctc	2580
cgcttcacaga	caagctgtga	ccgtctccgg	gagctgagc	cgccaggggt	tcctcactgt	2640
atcaccgaaa	cgcgcgaggg	agctgcggta	aagctccatc	ggctgctctg	gcagcgattc	2700
acagatgtct	gcctggttat	ccgcgtccag	ctcgctgagt	ctctccagaa	gcgttaattg	2760
ctggctctct	ataaagcggg	ccatgttaag	ggcggtctct	ctctgcttgg	tcacttgatg	2820
ctctcgtgta	aggggggaatt	ctctgtccatg	gggttaatga	taacgatgaa	acgagagagg	2880
atgctcacga	taagggttao	tgatgatgaa	cattgcctgg	taactggaac	ctgtgagggg	2940
aaacaaactg	cgggtatggat	gcggcgggac	cagagaaaaa	tcactccagg	tcactgcacg	3000
cgaaacgcca	gcaagacgta	gcacagcgcg	tgggcgcgca	tgccggcgat	aatggcctcg	3060
ctcccgccga	aacttttggg	ggcgggacba	gtgaagagg	cttgagcgag	ggcgcgcaag	3120
atcccgaaata	cgcgaagcga	cagggccgac	atcgtcgcg	tcacagcgaaa	gcgctccctg	3180
cgaaaaaatga	ccacagagcg	cggcggcacc	ctgcttaoga	gctgcattgat	aaagaagaca	3240
gtcttaagtg	gggggaagat	agtgatggcc	ctcccccacc	ggaaggagct	gaatgggggg	3300
aaggctctca	agggcatcgg	tcagatcccc	ggtgcctaat	gagtcgagca	acttaactta	3360
attcggtctg	gctcactgcc	cgcttccacg	tggggaaaac	cgtcgtgcba	gttcacttaa	3420
tgaatcggtg	caaggcgctc	taagagccat	ctcgctattg	ggcgccaggg	tgttctctcc	3480
ctccacccga	gagaggggta	ataactgatt	gccttccacc	gctggccctc	gaagagagcg	3540
cagcaagcgg	tcacagctgg	ctctgcacac	caggcgaaaa	ctcgtcttga	tggtggccaa	3600
ggggggggtg	taactctgag	ctctgtccat	atctgcttat	ccactaccg	agatatcccg	3660
accaaagcgc	actgcggaat	cgttaatggc	gcgcattcgg	ccagagcgca	ctcgatcgct	3720
ctctctctct	ctctgagctg	ctctctctct	ctcattccag	ctctgagctg	ctctctgaaa	3780
acgggacatg	gccttccagc	cgcttccgg	ctctgctctc	ctctgagctg	ctctctgaaa	3840
ctctctctct	ctctctctct	ctctctctct	ctctctctct	ctctctctct	ctctctctct	3900
caacagcgcg	attctgctgg	gacccatttg	gactaatat	ctctctctct	ctctctctct	3960
gactctctct	gacccatttg	gactctctct	gactctctct	ctctctctct	ctctctctct	4020
cgccgggaca	ttagtgcagc	cagcttccac	agcaatggca	ctctggctat	ctctctctct	4080
gttaactgct	agcccaactga	cgcgttgcac	caaaaatttg	tgcaacgcgc	ctctcagggc	4140
ttcgacgcgc	cttctgttata	ccatccacac	caccacgcgc	gcacccagct	gtctggcgcg	4200
agatttaata	gcgcgcacaa	tttgcgcagc	cgctgcagcg	gcacagctgg	aggtggccaa	4260
gcacatccagc	aacgactggt	tcgccgcag	ctggttctgc	acgcggttgg	gaatgtaatt	4320
ctctctctct	ctctctctct	ctctctctct	ctctctctct	ctctctctct	ctctctctct	4380
cttccacccc	ccggaaaacg	cttgataaga	gaccccgga	ctctctctct	ctctctctct	4440

```

cyttaactggt ttacatttca ccacccctgaa ttgaactctct tccggggcgct atcatgcat 4500
acccggaaaag gttttgggccc attcgatggt gtcccgagtc tggagctct cccttatgag 4500
actccatgat taggaagcag ccagtagta ggttgaggcc gttgagcag ccggccggcaa 4500
ggaatgggtgc atgcccgccl ttggtcttca agtatttatt cccaattcca ggcataaat 4500
aaaaacgaaa gctcagtcga aagaactggc ctttctgttt atctgttgtt tgtcgggtga 4500
cgctctctct agtaggacaa atccggccgg agcgatttg aacgttgoga agcaacggcc 4500
cggaggggtg cgggagggac gcccgccata aactggcagg aattaattcc aggcatacaa 4500
taaaaacgaaa ggtcagtcg aaagaactgg ccttctgttt tatctgttgt ttgtcgggtga 4500
acgtctctct gaggaggaca aatccggccg gagcgatttt gaacgttgog aagcaacggc 4500
ccggaggggtg cggggcagga ccggccgcac aaaactggcag gaattaattc caggcatcaa 5000
ataaaaacgaa aggtcagtc gaaagactgg gcttctgttt ctatctgttg ttgtcgggtg 5000
aacgtctctc tgagtaggac aaatccggcc ggagcgattt tgaacgttgc gaagcaacgg 5000
ccggaggggtg gggggcagg agcccgccca taaaactgca ggaatttatt ccaggcatca 5000
ataaaaacgaa aaggctcagt cgaagactg ggcttctgtt ttatctgttg gttgtcgggt 5000
gaaactctct ctgagttaga caaatccggc gggagcggat ttgaacgttg cgaagcaacg 5000
gcccgagggtg tggggggcag gaccccgccc ataaaactgc aggaattggg gatcgggaat 5000
aattccgggt ttaaacggg gatctcgatc ccggcaaat aatacgaact acctataggg 5000
aattgtgagc ggataacat tccctcttag aaataatttt gcttaacttt aagaaggaga 5000
tatacatatg gctagctcgc gagtogaagg ccggccggaa ttctcgagg gctcttctct 5000
catcaccggga gatgcactag ttgcccatac ccgagggcag tgggtacgca tgcgcgacat 5000
cgtgcggggt ggcgggcaca acagtgcaca cgcacatgac ctgaaagtcc ttgacgggca 5000
tggcaatccc gtagctcggc accggctgtt caactccggc gagcatccgg tgtacacggt 5000
ggttaggtg gaaggtctgc ctatgcacgg caccgggaac caccggttgt tgtgttttgt 5000
cgactgcgc ggggtgcoga cctgctgtgt gaagctgac gacgaaatca agccggggga 5000
taacggggtg atttaacgca ggcattcag ggtgagctg gtaggtttcc ccggggggaa 5000
accggaattt gggcccaaaa ctatcacagt cgggtcttcc ggaactgggt gttctctgtt 5000
agcacaccac cagagaccgg accgccaagg tctcgagac gagctgacgg accgggggtt 5000
ctactacggc aaagtgcaca ggtcacggc ccgggggtgt cagccgggtg atagcccttg 5000
tgctgacacg gcagaccacg cgtttatcac gaacgggttc gtcagccacg ctactgggtt 5000
cacgggtctg aactcagggc tcacgacaaa tctgtgtgta tccgcttggc aggtcaacac 5000
agcttatact gggggacaaat tggtcacata taacggcag acgtataaat gtttgcagcc 5000
ccacacctcc ttggcaggat gggaaacata caacttctt gccttctgtg agcttcaatg 5000
actgcaggaa ggggatccgg ctgctaacaa agcccgaag gaagctgagt tggctgtgtg 5000
caccgctgag caataactag cataacccct tggggcctct aaacgggtct tgaggggttt 5000
tttgtgaaa ggaggaaacta tatccggat 5000

```

(210) 6

(211) 30

(212) PRT

(213) Artificial Sequence

(220)

(223) Description of Artificial Sequence: synthetic peptide

422

Lys Ala Tyr Lys Thr Thr Glu Ala Asn Lys His Ile Ile Val Ala Tyr

Glu Gly Asn Pro Tyr Val Pro Val His Ile Asp Ala Thr Val

20

25

30

42100: 7
42110: 13
42120: FRT
42130: Artificial Sequence

42200:
42230: Description of Artificial Sequence: the amino acid
sequence deduced from the polylinker region of
pTXB1

44000: 7
Met Ala Met Gly Gly Gly Arg Leu Glu Gly Ser Ser Cys
1 5 10

42100: 8
42110: 42
42120: INA
42130: Artificial Sequence

42200:
42230: Description of Artificial Sequence: polylinker
region upstream of the modified intein from the
gyrA gene of Mycobacterium xenopi in pTXB1

43000: 8
ataaggaca tgggtgggga ccgcctcgag ggcctcttct gc 42

42100: 9
42110: 7
42120: FRT
42130: Artificial Sequence

42200:
42230: Description of Artificial Sequence: synthetic
peptide

44000: 9
Cys Asp Pro Glu Lys Asp Ser
1 5